I CLAIM:

1. A method of voice recognition, comprising the steps of:

organizing a plurality of speaker data points, representing a plurality of enrollment speakers, into a data structure using high-dimensional vectors that represent characteristics of enrollment voice samples from the enrollment speakers;

estimating a density of a subset of the plurality of speaker data points comprising the approximate nearest neighbors to an unidentified voice sample from an unidentified speaker; and

identifying the unidentified speaker based on one or more speaker data points most closely matching the unidentified voice sample as indicated by the estimated density.

2. The method of claim 1, wherein the step of estimating the density comprises estimating a probability density function using Parzen windows to estimate the probability density function.

3. The method of claim 1, wherein the step of estimating the density comprises estimating the density based on a distance between individual speaker data points within the subset of speaker data points

4. The method of claim 1, wherein the step of estimating the density further comprises controlling the relative contributions of individual speaker data points within the subset of speaker data points to the density based on a distance to a speaker data point from the unidentified voice sample.

5. The method of claim 1, wherein the step of estimating the density comprises estimating the density of the subset of speaker data points independent of parametric distribution information related to the plurality of speaker data points.

6. The method of claim 1, wherein the data structure module organizes the plurality of speaker data points such that a distance between individual speaker data points is based on characteristic similarities between associated voice samples, the distance measured in terms of one from the group containing: a Euclidean distance, a Minkowski distance, and a Manhattan distance.

7. The method of claim 1, wherein the data structure comprises a kd-tree.

8. The method of claim 1, wherein the plurality of speaker data points comprises a relatively large number of speaker data points.

9. The method of claim 1, further comprising a step of retrieving the subset of speaker data points using an unidentified speaker data point from the unidentified voice sample as an index into the plurality of speaker data points.

10. The method of claim 9, wherein the step of retrieving the subset of speaker data points comprises retrieving approximate nearest neighbors to the unidentified speaker data point, the approximate nearest neighbors comprising speaker data points within a distance calculated as a function of a distance of an absolute nearest neighbor.

11. The method of claim 1, wherein the subset of speaker data points includes more than one speaker data points associated with a common identification, and the step of identifying the unidentified speaker accumulates a score for the common identification.

12. The method of claim 1, further comprising extracting the high-dimensional vectors from the enrollment voice samples and the unidentified voice sample.

13. The method of claim 1, wherein the step of identifying the unidentified speaker comprises identifying the unidentified speaker as one of the enrollment speakers if matching is within an error threshold.

14. The method of claim 1, wherein an enrollment voice sample and the unidentified voice sample of a common speaker are text-independent.

15. A method of voice recognition, comprising the steps of:

retrieving a subset of speaker data points by using an unidentified speaker data point as an index into a data structure comprising a plurality of speaker data points, the subset of speaker data points representing approximate nearest neighbors to the unidentified speaker data;

estimating a probability density function from a subset of the plurality of speaker data points; and

identifying the unidentified speaker based on one or more speaker data points most closely matching the unidentified voice sample as indicated by the probability density function.

16. The method of claim 15, wherein the step of estimating the probability density function comprises estimating the probability density function using Parzen windows to estimate the probability density function.

17. A voice recognition system, comprising:

means for organizing a plurality of speaker data points, representing a plurality of enrollment speakers, into a data structure using high-dimensional vectors that represent characteristics of enrollment voice samples from enrollment speakers;

means for estimating a density of a subset of the plurality of speaker data points comprising the approximate nearest neighbors to an unidentified voice sample from an unidentified speaker; and

means for identifying the unidentified speaker based on one or more speaker data points most closely matching the unidentified voice sample as indicated by the estimated density.

18. The system of claim 17, wherein the means for estimating uses Parzen windows to estimate the density.

19. The system of claim 17, wherein the means for estimating estimates the density based on a distance between individual speaker data points within the subset of speaker data points.

20. The system of claim 17, wherein the means for estimating includes a smoothing parameter to control the relative contributions of individual speaker data points within the subset of speaker data points to the probability density function based on a distance to a speaker data point from the unidentified voice sample.

21. The system of claim 17, wherein the means for estimating estimates the density of the subset of speaker data points independent of parametric distribution information related to the plurality of speaker data points.

22. The system of claim 17, wherein the means for organizing organizes the plurality of speaker data points such that a distance between individual speaker data points is based on characteristic similarities between associated voice samples, the distance measured in terms of one from the group containing: a Euclidean distance, a Minkowski distance, and a Manhattan distance.

23. The system of claim 17, wherein the means for organizing comprises a kd-tree.

24. The system of claim 17, wherein the plurality of speaker data points comprises a relatively large number of speaker data points.

25. The system of claim 17, further comprising means for retrieving the subset of speaker data points uses an unidentified speaker data point from the unidentified voice sample as an index into the plurality of speaker data points.

26. The system of claim 25, wherein the means for retrieving the subset of speaker data points retrieves approximate nearest neighbors to the unidentified speaker data point, the approximate nearest neighbors comprising speaker data points within a distance calculated as a function of a distance of an absolute nearest neighbor.

27. The system of claim 17, wherein the subset of speaker data points includes more than one speaker data points associated with a common identification, and the identification module accumulates a score for the common identification.

28. The system of claim 17, further comprising a means for extracting the high-dimensional vectors from voice samples.

29. The system of claim 17, wherein the means for identifying identifies the unidentified speaker as one of the enrollment speakers if matching is within an error threshold.

30. The system of claim 17, wherein an enrollment voice sample and the unidentified voice sample of a common speaker are text-independent.

31. A computer program product, comprising:

a computer-readable medium having computer program instructions and data embodied thereon for voice recognition, comprising the steps of:

organizing a plurality of speaker data points, representing a plurality of enrollment speakers, into a data structure using high-dimensional vectors that represent characteristics of enrollment voice samples from the enrollment speakers;

estimating a density of a subset of the plurality of speaker data points comprising the approximate nearest neighbors to an unidentified voice sample from an unidentified speaker; and

identifying the unidentified speaker based on one or more speaker data points most closely matching the unidentified voice sample as indicated by the estimated density.

32. The computer program product of claim 31, wherein the step of estimating the density comprises estimating a probability density function using Parzen windows to estimate the probability density function.

33. The computer program product of claim 31, wherein the step of estimating the density comprises estimating the density based on a distance between individual speaker data points within the subset of speaker data points

34. The computer program product of claim 31, wherein the step of estimating the density further comprises controlling the relative contributions of individual speaker data points within the subset of speaker data points to the probability density function based on a distance to a speaker data point from the unidentified voice sample.

35. The computer program product of claim 31, wherein the step of estimating the density comprises estimating the probability density function of the subset of speaker data points independent of parametric distribution information related to the plurality of speaker data points.

36. The computer program product of claim 31, wherein the data structure module organizes the plurality of speaker data points such that a distance between individual speaker data points is based on characteristic similarities between associated voice samples, the distance measured in terms of one from the group containing: a Euclidean distance, a Minkowski distance, and a Manhattan distance.

37. The computer program product of claim 31, wherein the data structure comprises a kd-tree.

38. The computer program product of claim 31, wherein the plurality of speaker data points comprises a relatively large number of speaker data points.

39. The computer program product of claim 31, further comprising a step of retrieving the subset of speaker data points using an unidentified speaker data point from the unidentified voice sample as an index into the plurality of speaker data points.

40. The computer program product of claim 39, wherein the step of retrieving the subset of speaker data points comprises retrieving approximate nearest neighbors to the unidentified speaker data point, the approximate nearest neighbors comprising speaker data points within a distance calculated as a function of a distance of an absolute nearest neighbor.

41. The computer program product of claim 31, wherein the subset of speaker data points includes more than one speaker data points associated with a common identification, and the identification module accumulates a score for the common identification.

42. The computer program product of claim 31, further comprising extracting the high-dimensional vectors from the enrollment voice samples and the unidentified voice sample.

43. The computer program product of claim 31, wherein the step of identifying the unidentified speaker comprises identifying the unidentified speaker as one of the enrollment speakers if matching is within an error threshold.

44. The computer program product of claim 31, wherein an enrollment voice sample and the unidentified voice sample of a common speaker are text-independent.